AN IMPROVEMENT UNDER CERTAIN CIRCUMSTANCES OVER S. S. GUPTA'S PROCEDURE FOR SELECTING A SUBSET CONTAINING THE BEST NORMAL POPULATION

by

Arthur Roth

Technical Report No. 248

April, 1975

University of Minnesota
Minneapolis, Minnesota

1. Introduction.

Suppose we have k normal populations with known common variance σ^2, and we take random samples of size n (fixed and pre-specified) from each of them. We let $\mu_{(i)}$ $(i = 1,2,\ldots,k)$ be the i^{th} smallest population mean, and we assume $\mu_{(k)} > \mu_{(k-1)}$. The population whose mean is $\mu_{(k)}$ is called the best population. We wish to devise a procedure for selecting a subset of these populations such that the probability that this subset contains the best population is at least P^*, where $\frac{1}{k} < P^* < 1$ is also pre-specified. (This is called the P^*-condition.) Of course, good procedures are defined to be those with small expected subset size. Let $\bar{X}_i$ be the i^{th} sample mean, the largest of which is $\bar{X}_{max}$. Then Gupta [2] derives a procedure which includes the i^{th} population in the selected subset if and only if

$$(1.1) \qquad \bar{X}_i \geq \bar{X}_{max} - D^*$$

where D^* is defined by

$$(1.2) \qquad \int_{-\infty}^{\infty} \Phi^{k-1}\left(x + \frac{\sqrt{n}\, D^*}{\sigma}\right) d\Phi(x) = P^*.$$

D^* can be determined easily from (1.2) since the numbers $\tau(k,P^*)$ defined by

$$(1.3) \qquad \int_{-\infty}^{\infty} \Phi^{k-1}(x + \tau(k,P^*))\, d\Phi(x) = P^*$$

are positive for $P^* > \frac{1}{k}$ and tabled extensively in [1]. Gupta shows that his procedure satisfies the P^*-condition <u>for any configuration of the $\mu_{(i)}$'s,</u> even if $\Delta = \mu_{(k)} - \mu_{(k-1)}$ is arbitrarily small. Furthermore, D^* is the smallest value for which this is true.

1. Introduction.

Suppose we have k normal populations with known common variance σ^2, and we take random samples of size n (fixed and pre-specified) from each of them. We let $\mu_{(i)}$ $(i = 1,2,\ldots,k)$ be the i^{th} smallest population mean, and we assume $\mu_{(k)} > \mu_{(k-1)}$. The population whose mean is $\mu_{(k)}$ is called the best population. We wish to devise a procedure for selecting a subset of these populations such that the probability that this subset contains the best population is at least P^*, where $\frac{1}{k} < P^* < 1$ is also pre-specified. (This is called the P^*-condition.) Of course, good procedures are defined to be those with small expected subset size. Let $\bar{X}_i$ be the i^{th} sample mean, the largest of which is $\bar{X}_{max}$. Then Gupta [2] defines a procedure which includes the i^{th} population in the selected subset if and only if

$$\bar{X}_i \geq \bar{X}_{max} - D^* \qquad (1.1)$$

where D^* is defined by

$$\int_{-\infty}^{\infty} F^{k-1}\left(x + \frac{\sqrt{n}}{\sigma} D^*\right) dF(x) = P^* . \qquad (1.2)$$

D^* can be determined easily from (1.2) since the numbers $\lambda(k,P^*)$ defined by

$$\int_{-\infty}^{\infty} F^{k-1}(x + \lambda(k,P^*))\, dF(x) = P^* \qquad (1.3)$$

are positive for $P^* > \frac{1}{k}$ and tabled extensively in [1]. Gupta shows that his procedure satisfies the P^*-condition for any configuration of the $\mu_{(i)}$'s, even if $\mu_{(k)} - \mu_{(k-1)}$ is sufficiently small. Further, P^* is the smallest value for which this is true.

- 1 -

The idea of an indifference zone is usually incorporated into ranking and selection problems of this type. That is, we are indifferent as to whether or not the P^*-condition is satisfied if $\Delta < d^*$, where $d^* > 0$ is usually pre-specified. I will NOT pre-specify d^*. Instead, I will introduce a procedure, which I call the fiducial procedure, that is independent of d^*. I then show that (in terms of subset size) this procedure is uniformly better than Gupta's procedure but that it fails to satisfy the P^*-condition for small values of Δ except when $k = 2$. I then find a value d_0^* of Δ which is sufficiently large to guarantee that the fiducial procedure satisfies the P^*-condition. If we are willing to define an indifference zone by $\Delta < d^*$, then we have improved upon Gupta's procedure whenever $d^* \geq d_0^*$, as well as for some smaller (but we don't know how small) values of d^*.

Finally, we note that the introduction of d^* allows us to use Gupta's procedure with a smaller value of D^* than the one defined by (1.2) and still satisfy the P^*-condition. If we modify Gupta's procedure in this way, the fiducial procedure is no longer uniformly better than it. However, this modification of Gupta's procedure requires us to incorporate d^* into the actual procedure. In contrast, the fiducial procedure still has the advantage of being independent of d^*.

2. Definition of the Fiducial Procedure.

If we are willing to treat the population means μ_i, which are unknown, as fiducial random variables based on the observations, then μ_i has a fiducial distribution which is normal with mean $\bar{X}_i$ and variance σ^2/n.

The idea of an indifference zone is usually incorporated into ranking and selection problems of this type. That is, we are indifferent as to whether or not the P^*-condition is satisfied if $|\delta| < d^*$, where $d^* > 0$ is usually pre-specified. I will NOT pre-specify d^*. Instead, I will introduce a procedure, which I call the Fiducial procedure, that is independent of d^*. I then show that (in cases of interest mine) this procedure is uniformly better than Gupta's procedure but that it fails to satisfy the P^*-condition for small values of δ, except when $k = 2$. I then find a value d_0^* of d^* which is sufficiently large to guarantee that the Fiducial procedure satisfies the P^*-condition. If we are willing to define an indifference zone by $|\delta| < d^*$, then we have improved upon Gupta's procedure whenever $d^* \geq d_0^*$, as well as for some smaller (but we don't know how small) values of d^*.

Finally, we note that the introduction of d^* allows us to use Gupta's procedure with a smaller value of d than the one defined by (1.2) and still satisfy the P^*-condition. If we modify Gupta's procedure in this way, the Fiducial procedure is no longer uniformly better than it. However, this modification of Gupta's procedure requires us to incorporate d^* into the actual procedure. In contrast, the Fiducial procedure still has the advantage of being independent of d^*.

2. Definition of the Fiducial Procedure.

If we are willing to treat the population means μ_i, which are unknown, as fiducial random variables based on the observations, then μ_i has a fiducial distribution which is normal with mean $\bar{X}_i$ and variance σ^2/n.

These "random variables" are also independent. Let P_i be the "probability" that μ_i will be the largest if we observe these random variables. Then

$$
\begin{aligned}
(2.1) \qquad P_i &= P\{\mu_j < \mu_i, \ 1 \le j \le k, \ j \neq i\} \\
&= P\{\frac{\mu_j - \bar{X}_j}{\sigma/\sqrt{n}} < \frac{\mu_i - \bar{X}_i}{\sigma/\sqrt{n}} + \frac{\sqrt{n}\,(\bar{X}_i - \bar{X}_j)}{\sigma}, \ 1 \le j \le k, \ j \neq i\} \\
&= \int_{-\infty}^{\infty} \prod_{\substack{j=1 \\ j \neq i}}^{k} \Phi(x + \frac{\sqrt{n}\,(\bar{X}_i - \bar{X}_j)}{\sigma})\, d\Phi(x) \quad (i = 1,2,\ldots,k) .
\end{aligned}
$$

We first choose the population with the largest P-value (equivalently, the largest sample mean), then the second largest, etc., and we continue until we have included enough populations that their summed P-values are not less than P^*. (Of course, $\sum P_i = 1$, so this can always be done.) This completes the definition of the procedure.

3. Comparison of the Fiducial Procedure to Gupta's Procedure.

Let S_F and S_G denote the random subsets chosen, respectively, by the fiducial and the Gupta procedures.

<u>LEMMA 3.1:</u> $S_F \subseteq S_G$ no matter what the observed sample means are.

<u>PROOF</u>: Suppose $j \notin S_G$. Then there is an integer $i \neq j$ such that $\bar{X}_j \le \bar{X}_i - D^*$.

<u>Case 1:</u> Suppose $\bar{X}_j \ge \bar{X}_m$ for every $m \neq i,j$. Then $\bar{X}_i - \bar{X}_r \ge D^*$ for all $r \neq i$. From (2.1), we see that

$$
(3.1) \qquad P_i \ge \int_{-\infty}^{\infty} \Phi^{k-1}(x + \frac{\sqrt{n}\, D^*}{\sigma})\, d\Phi(x) = P^* ,
$$

the last equality coming from (1.2), the definition of D^*. Hence, the i^{th} population is the only one chosen, so that $S_F = S_G = \{i\}$. In particular, $j \notin S_F$.

These "random variables" are also independent. Let P_i be the "probability" that μ_i will be the largest if we observe these random variables. Then

$$(2.1) \qquad P_i = P\{Y_j < Y_i,\ j = 1,\ldots,k,\ j \neq i\}$$

$$\equiv P\left\{\frac{Y_j - \bar{X}_j}{\sigma/\sqrt{n}} < \frac{Y_i - \bar{X}_i}{\sigma/\sqrt{n}} + \frac{\sqrt{n}\,(\bar{X}_i - \bar{X}_j)}{\sigma},\ j = 1,\ldots,k,\ j \neq i\right\}$$

$$= \int_{-\infty}^{\infty} \prod_{\substack{j=1 \\ j\neq i}}^{k} \Phi\left(x + \frac{\sqrt{n}\,(\bar{X}_i - \bar{X}_j)}{\sigma}\right) d\Phi(x) \quad (i = 1,2,\ldots,k).$$

We first choose the population with the largest T-value (equivalently, the largest sample mean), then the second largest, etc., and we continue until we have included enough populations that their summed P-values are not less than P^*. (Of course, $\sum P_i = 1$, so this can always be done.) This completes the definition of the procedure.

3. <u>Comparison of the Fiducial Procedure to Gupta's Procedure.</u>

Let S_F and S_G denote the random subsets chosen, respectively, by the fiducial and the Gupta procedures.

<u>LEMMA 3.1</u>: $S_F \subseteq S_G$ no matter what the observed sample means are.

<u>PROOF</u>: Suppose $j \notin S_G$. Then there is an integer $i \neq j$ such that $\bar{X}_j < \bar{X}_i - D^*$.

<u>Case 1</u>: Suppose $\bar{X}_u \leq \bar{X}_j$ for every $u \neq i,j$. Then $\bar{X}_i - \bar{X}_r > D^*$ for all $r \neq i$. From (2.1), we see that

$$(3.1) \qquad P_i \geq \int_{-\infty}^{\infty} \Phi^{k-1}\left(x + \frac{D^*\sqrt{n}}{\sigma}\right) d\Phi(x) = P^*,$$

the last equality coming from (1.2), the definition of D^*. Hence, the i^{th} population is the only one chosen, so that $S_F = S_G = \{i\}$. In particular, $j \notin S_F$.

Case 2: Let $s \geq 1$ population means lie between $\bar{X}_j$ and $\bar{X}_i$, i.e., all other means $\leq \bar{X}_j \leq \bar{X}_{i_1} \leq \bar{X}_{i_2} \leq \ldots \leq \bar{X}_{i_s} \leq \bar{X}_i$.

Let $\{P_m\}_{m=1,2,\ldots,k}$ denote the appropriate P-values for this situation. Let $\{P_m^*\}_{m=1,2,\ldots,k}$ denote the appropriate P-values for the situation where we modify the configuration by removing all population means between $\bar{X}_i$ and $\bar{X}_j$ and placing them to the left of $\bar{X}_j$, i.e., replace $\bar{X}_{i_m}$ by $\bar{X}_j - \epsilon$ for some $\epsilon > 0$ and all $m = 1,2,\ldots,s$.

From Case 1, we see that $P_i^* \geq P^*$.

From (2.1), it is clear that P_m is an increasing function of $\bar{X}_m$ and a decreasing function of $\bar{X}_r$ for all $r \neq m$. Furthermore, all the P_m's remain the same if each $\bar{X}_m$ is increased or decreased by the same constant.

Let $T = \{i_1, i_2, \ldots, i_s\}$. It follows that $P_m^* < P_m$ if and only if $m \in T$.

Now $\sum P_i = \sum P_i^* = 1$; hence

$$(3.2) \qquad 0 = \sum (P_m^* - P_m) = \sum_{m \notin T} (P_m^* - P_m) + \sum_{m \in T} (P_m^* - P_m)$$

$$\geq P_i^* - P_i + \sum_{m \in T} (P_m^* - P_m)$$

which means that $P_i + \sum_{m \in T} P_m \geq P_i^* + \sum_{m \in T} P_m^* > P_i^* \geq P^*$. Since $P_i, P_{i_1}, P_{i_2}, \ldots, P_{i_s}$ are the $(s+1)$ largest P-values, we see that $S_F \subseteq \{i\} \cup T$. In particular, $j \notin S_F$.

LEMMA 3.2: $P(S_F = S_G) < 1$ (i.e., there is positive probability that S_F is a proper subset of S_G) whenever $k > 2$.

PROOF: In Case 2 of Lemma 3.1, we proved that $P_i + P_{i_1} + \ldots + P_{i_s}$ was strictly greater than P^* even when we took $\bar{X}_j = \bar{X}_i - D^*$ (as long as $s \geq 1$). Now $P_i + P_{i_1} + \ldots + P_{i_s}$ is a decreasing continuous function of

CASE 2: Let a m.l. population means lie between $\bar{\mu}_j$ and $\bar{\mu}_i$, i.e., all other means $\bar{\mu}_j < \bar{\mu}_{i_1} < \bar{\mu}_{i_2} < \ldots < \bar{\mu}_{i_s} < \bar{\mu}_i$.

Let $(P_r)_{r=1,2,\ldots,k}$ denote the appropriate P-values for this situation. Let $(P_r^*)_{r=1,2,\ldots,k}$ denote the appropriate P-values for the situation where we modify the configuration by removing all population means between $\bar{\mu}_j$ and $\bar{\mu}_i$ and placing them to the left of $\bar{\mu}_j$, i.e., replace $\bar{\mu}_{i_m}$ by $\bar{\mu}_j - \epsilon_m$ for some $\epsilon_m > 0$ and all $m = 1,2,\ldots,s$.

From Case 1, we see that $P_i^* \geq P^*$.

From (2.1), it is clear that P_m is an increasing function of $\bar{\mu}_m$ and a decreasing function of $\bar{\mu}_r$ for all $r \neq m$. Furthermore, all the P_m's remain the same if each $\bar{\mu}_r$ is increased or decreased by the same constant. Let $T = \{i_1, i_2, \ldots, i_s\}$. It follows that $P_m^* < P_m$ if and only if $m \in T$. Now $\sum P_r = \sum P_r^* = 1$; hence

$$(4.2) \qquad 0 = \sum (P_r^* - P_r) = \sum_{m \notin T} (P_m^* - P_m) + \sum_{m \in T} (P_m^* - P_m)$$

$$\geq P_i^* - P_i + \sum_{m \in T} (P_m^* - P_m)$$

which means that $P_i + \sum_{m \in T} P_m \geq P_i^* + \sum_{m \in T} P_m^* > P_i^* \geq P^*$. Since $P_i, P_{i_1}, P_{i_2}, \ldots, P_{i_s}$ are the $(s+1)$ largest P-values, we can find $S_R \subseteq \{i\} \cup T$. In particular, $j \notin S_R$.

LEMMA 4.2: $P(S_R = S_G) < 1$ (i.e., there is positive probability that S_R is a proper subset of S_G) whenever $k > 2$.

PROOF: In Case 2 of Lemma 4.1, we proved that $P_i + P_{i_1} + \ldots + P_{i_s}$ was strictly greater than P^* even when we took $\bar{\mu}_i = \bar{\mu}_j - \epsilon^*$ (as long as $s \geq 1$). Now $P_i + P_{i_1} + \ldots + P_{i_s}$ is a decreasing continuous function of

$\bar{X}_j$. Furthermore, since the integrand in the last expression in (2.1) is bounded, we can differentiate inside this integral for each P_m to show that the partial derivative of this function with respect to $\bar{X}_j$ is finite. Hence, there is a number $\epsilon > 0$ such that we can increase $\bar{X}_j$ to $\bar{X}_j + \epsilon$ without decreasing $P_i + P_{i_1} + \ldots + P_{i_s}$ below P^*. Hence, $S_F \subseteq \{i\} \cup T$ as before. But now $\bar{X}_i - \bar{X}_j = D^* - \epsilon < D^*$, so that $j \in S_G$ and $S_G \nsubseteq S_F$. This situation occurs whenever all of the following occur:

(i) $\bar{X}_i - \bar{X}_j \in (D^* - \epsilon, D^*)$

(ii) $\bar{X}_m \in (\bar{X}_i, \bar{X}_j)$ for all $m \in T = \{i_1, i_2, \ldots, i_s\}$

(iii) $\bar{X}_m < \bar{X}_j$ for $m \notin \{i,j\} \cup T$.

For any fixed values of $i, j, i_1, i_2, \ldots, i_s$, the probability that (i)-(iii) will occur is positive, establishing the lemma.

Note that lemma 3.1 fails for the case $k = 2$ because the set T is then empty with probability one (since there can be no population means between the only two that we observe). This is very important, since the two corollaries that follow also fail. In particular, the P^*-condition is satisfied by the fiducial procedure for $k = 2$, as we shall see later.

COROLLARY 3.2.1: $E(\# S_F) < E(\# S_G)$ whenever $k > 2$. The proof is obvious.

COROLLARY 3.2.2: The fiducial procedure does not satisfy the P^*-condition when $k > 2$.

PROOF: The worst case clearly approaches the situation where the "best" population is arbitrarily specified beforehand and all the μ_i's are equal. In this case, all the $\bar{X}_i$'s are exchangeable (in fact, they are i.i.d.) so that

$$(3.3) \qquad P(\text{best population is in subset} \mid \text{the order statistic of the } \bar{X}_i\text{'s}) = \frac{\# S}{k}$$

T_i. Furthermore, since the integrand in the last expression in (2.1) is bounded, we can differentiate inside this integral for each T_i to show that the partial derivative of this function with respect to T_i is finite. Hence, there is a number $\epsilon > 0$ such that we can increase T_i to $T_i + \epsilon$ without decreasing $T_1 + T_2 + \dots + T_{i_s}$ below T^*. Hence, $S_k \in (i) \cup T$ as before. But now $T_i - T_i' = D^* - \epsilon < D^*$, so that $i \in I_G$ and $I_G \neq I_k$. This situation occurs whenever all of the following occur:

(i) $T_i - T_i' \in (D^* - \epsilon, D^*)$

(ii) $T_m \in (T_i, T_i')$ for all $m \in I = \{i_1, i_2, \dots, i_s\}$

(iii) $T_m < T_i$ for $m \notin \{i, i_1\} \cup T$.

For any fixed values of $i, i', i_1, i_2, \dots, i_s$, the probability that (i)-(iii) will occur is positive, establishing the lemma.

Note that Lemma 3.1 fails for the case $k = 2$ because the set W is then empty with probability one (since there can be no population means between the only two that are observed). This has important implications, since the two corollaries that follow also fail. In particular, the P^*-condition is satisfied by the Bechhofer procedure for $k = 2$, as we shall see later.

COROLLARY 3.2.1: $E(\mu_{S_k}) < E(\mu_{S_G})$ whenever $k > 2$. The proof is obvious.

COROLLARY 3.2.2: The Bechhofer procedure does not satisfy the P^*-condition when $k > 2$.

PROOF: The worst case clearly approaches the situation where the "best" population is arbitrarily specified beforehand and all the μ_i's are equal. In this case, all the T_i's are exchangeable (in fact, they are i.i.d.) so that

(3.2) P(best population is included | the order statistics of the T_i's)

$$= \frac{s}{k}$$

independently of the $\bar{X}_i$'s where $\#S$ is the size of the selected subset S. Hence, unconditionally (letting CS stand for "correct selection"), we have

$$(3.4) \qquad P(CS) = E[P(CS)|\text{the order statistic}] = \frac{1}{k} E(\# S) \;.$$

But $\frac{1}{k} E(\# S_F) < \frac{1}{k} E(\# S_G)$, and (in this worst case) this last quantity is equal to P^*, as shown in [2].

4. Calculation of d_0^*.

For convenience of notation, assume the $k^{\underline{th}}$ population is best, i.e., $\mu_k = \mu_{(k)}$. Let $Y_j = \sqrt{n}\,(\bar{X}_k - \bar{X}_j)/\sigma$ for $j = 1,2,\ldots,k-1$. Then the Y_j's have a joint multivariate normal distribution with common variance 2, common correlation 1/2, and $EY_j = \sqrt{n}\,(\mu_k - \mu_j)/\sigma$. Suppose $\mu_k - \mu_j \geq d^*$ $(j = 1,2,\ldots,k-1)$. Then $EY_j \geq \sqrt{n}\, d^*/\sigma$ for every $j = 1,2,\ldots,k-1$. Let $c_j = EY_j - \sqrt{n}\, d^*/\sigma > 0$, and let $V_j = Y_j - c_j$. Then the V_j's also are jointly multivariate normal with common variance 2, common correlation 1/2, and common mean $\sqrt{n}\, d^*/\sigma$. Furthermore $V_j \leq Y_j$ for each j. Now

$$(4.1) \qquad P_k = \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} \Phi(x + Y_j)\, d\Phi(x) \geq \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} \Phi(x + V_j)\, d\Phi(x) \;.$$

Since the functions $f_j(x) = \Phi(x + V_j)$ are all monotone increasing in x, we can write (see [3])

$$(4.2) \qquad P_k \geq \prod_{j=1}^{k-1} \int_{-\infty}^{\infty} \Phi(x + V_j)\, d\Phi(x) \;.$$

<u>LEMMA 4.1:</u> $\displaystyle\int_{-\infty}^{\infty} \Phi(ax+b)\,d\Phi(x) = \Phi\left(\frac{b}{\sqrt{1+a^2}}\right)$.

<u>PROOF:</u> Let Z_1 and Z_2 be i.i.d. standard normal. Then the left side of the above equation is equal to $P\{Z_1 \leq aZ_2 + b\} = P\{Z_1 - aZ_2 \leq b\}$ $= P\{(Z_1 - aZ_2)/\sqrt{1+a^2} \leq b/\sqrt{1+a^2}\} = \Phi(b/\sqrt{1+a^2})$.

Let $W_j = V_j/\sqrt{2}$. Then the W_j's are multivariate normal with common variance 1, common correlation 1/2, and common mean $\sqrt{n}\, d^*/\sigma\sqrt{2}$. Using (4.2) and lemma 4.1, we have

$$(4.3) \qquad P_k \geq \prod_{j=1}^{k-1} \Phi(W_j) \; .$$

Noting that we get a correct selection whenever $P_k > 1-P^*$, we will satisfy the P^*-condition if $P(P_k > 1-P^*) \geq P^*$. But

$$(4.4) \qquad P(P_k > 1-P^*) \geq P\left(\prod_{j=1}^{k-1} \Phi(W_j) > 1-P^*\right)$$

$$\geq P(\Phi(W_j) > (1-P^*)^{1/(k-1)} \quad \text{for } j = 1,2,\ldots,k-1)$$

$$= P(W_j > \Phi^{-1}[(1-P^*)^{1/(k-1)}] \quad \text{for } j = 1,2,\ldots,k-1)$$

$$= P(\sqrt{n}\, d^*/\sigma - \sqrt{2}\, W_j < \sqrt{n} d^*/\sigma - \sqrt{2}\, \Phi^{-1}[(1-P^*)^{1/(k-1)}]$$

$$\text{for } 1 \leq j \leq k-1) \; .$$

Now $\sqrt{n}\, d^*/\sigma - \sqrt{2}\, W_j$ are multivariate normal with common mean 0, common variance 2, and common correlation 1/2. If $Z_1, Z_2, \ldots, Z_k$ are i.i.d. standard normal, the random variables $R_i = Z_i - Z_k$ $(i = 1,2,\ldots,k-1)$ have the same distribution. Hence

$$(4.5) \quad P(P_k > 1-P^*) \geq P(Z_i - Z_k < \sqrt{n} d^*/\sigma - \sqrt{2}\, \Phi^{-1}[(1-P^*)^{1/(k-1)}]$$

$$\text{for } 1 \leq i \leq k-1)$$

$$= P(Z_i < Z_k + \sqrt{n} d^*/\sigma - \sqrt{2}\, \Phi^{-1}[(1-P^*)^{1/(k-1)}] \text{ for } 1 \leq i \leq k-1)$$

$$= \int_{-\infty}^{\infty} \Phi^{k-1}\left(x + \frac{\sqrt{n} d^*}{\sigma} - \sqrt{2}\, \Phi^{-1}[(1-P^*)^{1/(k-1)}]\right) d\Phi(x) \; .$$

We can now find d_0^* by equating $\sqrt{n} d^*/\sigma - \sqrt{2}\, \Phi^{-1}[(1-P^*)^{1/(k-1)}]$ to the value $\tau(k,P^*)$ defined in section 1.

It is interesting to compare d_0^* with Gupta's D^* even though these quantities have a different meaning. Since $\sqrt{n}\, D^*/\sigma = \tau(k,P^*)$, it is clear

that $d_0^* < D^*$ if and only if $\Phi^{-1}[(1-P^*)^{1/(k-1)}] < 0$, i.e., if and only if $1-P^* < (\frac{1}{2})^{k-1}$, which is satisfied for all small k-values up to a certain maximum. For example, if $P^* = .95$ then $d_0^* < D^* \Leftrightarrow k \le 5$. If $P^* = .99$, then $d_0^* < D^* \Leftrightarrow k \le 7$. We note that practical applications of ranking and selection problems usually do deal with small values of k. We again emphasize that d_0^* is not the smallest value of Δ which satisfies the P^*-condition; we used several rough inequalities just to obtain a value which we are sure will work.

It is worth noting that for the special case $k = 2$, we have $d_0^* = 0$. To show this, let $\tau = \tau(k,P^*)$ for convenience of notation. Then $P^* = \int_{-\infty}^{\infty} \Phi(x+\tau)d\Phi(x) = \Phi(\tau/\sqrt{2})$ by lemma (4.1). Hence $\tau/\sqrt{2} = \Phi^{-1}(P^*)$ so that $-\tau/\sqrt{2} = \Phi^{-1}(1-P^*)$. Hence, we find d_0^* by solving the equation $\sqrt{n}d^*/\sigma + \tau = \tau$; thus $d_0^* = 0$. The implication, of course, is that the fiducial procedure does satisfy the P^*-condition when $k = 2$. It is interesting and easy to verify that when $k = 2$, the fiducial and Gupta procedures are equivalent. Thus, even for $k = 2$, we have not found an unconditional improvement over Gupta's procedure. Furthermore, the fact that the P^*-condition is satisfied could have been obtained directly from this equivalence.

5. Modifying Gupta's Procedure if d^* exists.

In [2], the value D^* was obtained by deriving the fact that

$$(5.1) \quad P(CS) \ge \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} \Phi\left(x + \frac{\sqrt{n}D^*}{\sigma} + \frac{\sqrt{n}(\mu_{(k)} - \mu_{(j)})}{\sigma}\right)d\Phi(x)$$

and noting that the last term inside the Φ-function was non-negative. If d^* exists, however, this last term is then at least $\sqrt{n}d^*/\sigma$. Therefore, we can equate $\sqrt{n}(D^* + d^*)/\sigma$ [rather than just $\sqrt{n}D^*/\sigma$] to τ. Hence, D^* can be decreased by d^*. For large values of k, when we have seen that

that $d_0^* < d^*$ if and only if $[\ldots(1-P^*)^{1/(k-1)}] < 0$, i.e., if and only if $1-P^* < (\frac{1}{k})^{k-1}$, which is satisfied for all small k-values up to a certain maximum. For example, if $P^* = .95$ then $d_0^* < d^*$ for $k \le$ [illegible]. If $P^* = .99$, then $d_0^* < d^*$ for $k \le$ [illegible]. We note that practical applications of ranking and selection problems usually deal with small values of k. We again emphasize that d_0^* is not the smallest value of d which satisfies the P^*-condition; we used several rough inequalities just to obtain a value which we are sure will work.

It is worth noting that for the special case $k = 2$, we have $d_0^* = 0$. To show this, let $r = r(k,P^*)$ for convenience of notation. Then $P^* = \int_{-\infty}^{\infty} G(x+r)dG(x) = G(r/\sqrt{2})$ by Lemma (4.1). Hence $r/\sqrt{2} = G^{-1}(P^*)$ so that $r/\sqrt{2} = G^{-1}(P^*)$. Hence we find d_0^* by solving the equation $\sqrt{n}d^*/\sigma + r = r$; thus $d_0^* = 0$. The implication, of course, is that the Bechhofer procedure does satisfy the P^*-condition with $k = 2$. It is interesting and easy to verify that when $k = 2$, the Bechhofer and Gupta procedures are equivalent. Thus, even for $k = 2$, we have not found an unconditional improvement over Gupta's procedure. Furthermore, the fact that the P^*-condition is satisfied could have been obtained directly from this equivalence.

5. Modifying Gupta's Procedure if d^* exists.

In [2], the value d^* was obtained by making use of the fact that

$$(5.1)\quad E(\omega) = \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} G\left(x + \frac{\sqrt{n}\,d^*}{\sigma} + \ldots\right) dG(x)$$

and noting that the last term inside the G-function was non-negative. If d^* exists, however, this last term is then at least $\sqrt{n}d^*/\sigma$. Therefore, we can equate $\sqrt{n}(d + d^*)/\sigma$ [rather than just $\sqrt{n}d/\sigma$] to r. Hence, d can be decreased by d^*. For large values of k, when we have seen that

$D^* < d_0^*$, this must be uniformly better than the fiducial procedure since the subset size must then always be one. For small values of k, it is not clear which of the two procedures is better. Even for large k, however, the uniform superiority of Gupta's procedure only applies if we use d_0^*. But, as has been mentioned already, smaller values of d^* will work, too. Since we don't know just how small these values are, and since the modified Gupta procedure depends on the particular value of d^*, it is not clear which value we should use in the procedure. The fiducial procedure has no such difficulty. It is carried out indpendently of d^*, and it satisfies the P^*-condition as long as the indifference zone is at least as large as some minimum value (for which d_0^* is a very crude upper bound), which is unknown. Thus, the fiducial procedure has some advantages even relative to the modified Gupta procedure.

References

[1] Gupta, S. S., "Probability Integrals of Multivariate Normal and Multivariate t", AMS 34: 792-828 (1963).

[2] Gupta, S. S., "On Some Multiple Decision (Selection and Ranking) Rules", Technometrics 7: 225-245 (1965).

[3] Kimball, A. W. "On Dependent Tests of Significance in the Analysis of Variance", AMS 22: 600-602 (1951).